

# Building Trustworthy Generative Artificial Intelligence for Diabetes Care and Limb Preservation: A Medical Knowledge Extraction Case

Journal of Diabetes Science and Technology  
1–7

© 2024 Diabetes Technology Society


Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/19322968241253568

journals.sagepub.com/home/dst



**Shayan Mashatian, MS<sup>1,2</sup> , David G. Armstrong, DPM, MD, PhD<sup>3</sup>, Aaron Ritter, DPM<sup>4</sup>, Jeffery Robbins, DPM<sup>5</sup>, Shereen Aziz, BA<sup>2</sup>, Ilia Alenabi, HSD<sup>2</sup>, Michelle Huo, HSD<sup>6</sup>, Taneeka Anand, HSD<sup>7</sup>, and Kouhyar Tavakolian, PhD<sup>1</sup>**

## Abstract

**Background:** Large language models (LLMs) offer significant potential in medical information extraction but carry risks of generating incorrect information. This study aims to develop and validate a retriever-augmented generation (RAG) model that provides accurate medical knowledge about diabetes and diabetic foot care to laypersons with an eighth-grade literacy level. Improving health literacy through patient education is paramount to addressing the problem of limb loss in the diabetic population. In addition to affecting patient well-being through improved outcomes, improved physician well-being is an important outcome of a self-management model for patient health education.

**Methods:** We used an RAG architecture and built a question-and-answer artificial intelligence (AI) model to extract knowledge in response to questions pertaining to diabetes and diabetic foot care. We utilized GPT-4 by OpenAI, with Pinecone as a vector database. The NIH National Standards for Diabetes Self-Management Education served as the basis for our knowledge base. The model's outputs were validated through expert review against established guidelines and literature. Fifty-eight keywords were used to select 295 articles and the model was tested against 175 questions across topics.

**Results:** The study demonstrated that with appropriate content volume and few-shot learning prompts, the RAG model achieved 98% accuracy, confirming its capability to offer user-friendly and comprehensible medical information.

**Conclusion:** The RAG model represents a promising tool for delivering reliable medical knowledge to the public which can be used for self-education and self-management for diabetes, highlighting the importance of content validation and innovative prompt engineering in AI applications.

## Keywords

generative AI, knowledge extraction, LLM, diabetes self-education, layperson education, trustworthy AI, few-shot learning, diabetic foot ulcer

## Introduction

Since the wide availability of large language models (LLMs), they have sparked considerable interest among researchers across diverse disciplines owing to their remarkable text processing and generation capabilities.<sup>1</sup> Large language models have ushered in a new era of knowledge extraction and generation, with particular emphasis on their application in diverse contexts, including medical knowledge extraction.<sup>2</sup>

The existing body of research on this subject can be broadly categorized into two streams. The first stream focuses on extracting knowledge from general artificial intelligence (GAI) systems such as OpenAI's LLMs (eg,

<sup>1</sup>Biomedical Engineering Program, University of North Dakota, Grand Forks, ND, USA

<sup>2</sup>Silverberry Group, Inc., Grand Forks, ND, USA

<sup>3</sup>Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

<sup>4</sup>U.S. Department of Veterans Affairs, Little Rock, AR, USA

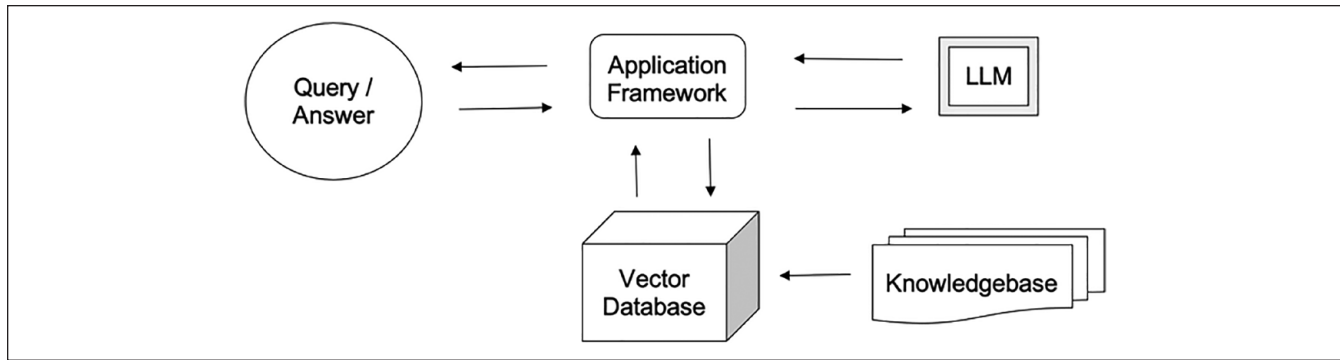
<sup>5</sup>U.S. Department of Veterans Affairs, Washington, DC, USA

<sup>6</sup>Department of Life Sciences, University of California, Los Angeles, Los Angeles, CA, USA

<sup>7</sup>Oxford College of Emory University, Oxford, GA, USA

## Corresponding Author:

Shayan Mashatian, MS, Biomedical Engineering Program, University of North Dakota, 4200 James Ray Drive, Grand Forks, ND 58202, USA.  
Email: shayan@silverberry.ai



**Figure 1.** Retriever-augmented generation (RAG) architecture model.

GPT-3 or GPT-4). This approach involves either the use of a public chat interface (ChatGPT) or the application of its Application Programming Interface (API). Researchers in this domain strive to craft optimal prompts to extract medical knowledge as a general domain, evaluate its efficacy, or concentrate on specific domains such as Cognitive Score.<sup>3</sup>

Simultaneously, a second stream of research has emerged, addressing the inherent challenges posed by the general nature of LLMs trained on unverified Internet content. The challenges posed by the general nature of LLMs necessitate the development of more controlled methods for knowledge extraction. One such promising method is retriever-augmented generation (RAG) (Figure 1), a technique in natural language processing and machine learning that combines the capabilities of two models: a retriever and a generator.<sup>4</sup> The retriever fetches relevant context or information from a large body of text, such as a database or the Internet, which is then provided to the generator, typically a LLM, to facilitate the generation of more accurate, relevant, and informed responses. The LLM can also be utilized to provide the answers in a conversational format.

### Problems Associated With Diabetic Foot Disease

Diabetic foot ulcers (DFUs) are a common, highly morbid consequence of longstanding and poorly managed diabetes.<sup>5</sup> Of the estimated 537 million people worldwide who have diabetes, one-third will develop DFU in their lifetime. Approximately 20% of people who develop a DFU will require lower-extremity amputation, either minor (below the ankle), major (above the ankle), or both,<sup>6</sup> and 10% will die within one year of their first DFU diagnosis.<sup>7,8</sup> Translating those numbers to the events in daily life means every second, someone with diabetes develops a foot ulcer, half of those wounds get infected, 20% of them end up in a hospital, and eventually, every 20 seconds, someone loses a limb to a DFU.<sup>9-12</sup>

Improving health literacy through patient education is paramount to addressing the problem of limb loss in the diabetic population. A patient who comprehends the

conversation or written material provided during a physician visit will forget 40% to 80% of the medical information discussed or distributed during that visit.<sup>13</sup> Novel approaches to education must be explored to improve health outcomes through patient education. Patient education that promotes health literacy through self-management education builds a partnership of accountability between physician and patient, enhancing communication and improving health outcomes.<sup>14</sup>

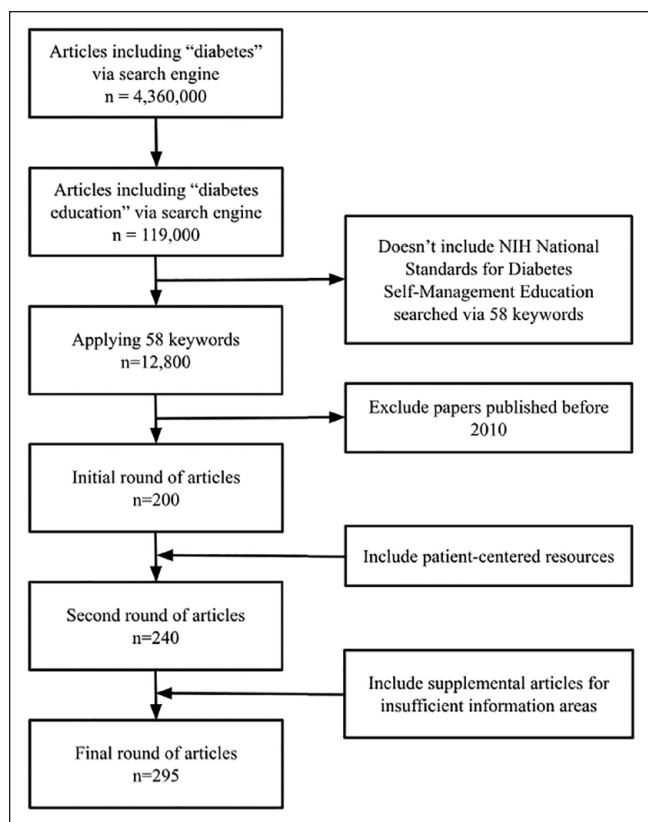
In addition to affecting patient well-being through improved outcomes, improved physician well-being is an important outcome of a self-management model for patient health education. Effective patient care promotes and requires physician well-being and improves health literacy.<sup>15</sup> As the patient experiences symptoms or has questions regarding lifestyle choices, a self-managed model provides trustworthy medical knowledge, reducing the administrative burden of the physician and aiding the physician in the education of the patient. Consistent patient education interventions improve patient satisfaction, adherence, and outcomes.<sup>16</sup> This novel use of a self-managed educational model to support health literacy and reduce the administrative burden directly promotes physician well-being.<sup>17</sup>

This study presents the implementation of a RAG model to utilize an AI model for knowledge extraction in response to questions pertaining to diabetes and diabetic foot care. The primary objective was to assess the system's ability to provide a layperson with an eighth-grade literacy level that is user-friendly, accurate, and comprehensible answers to medical questions.

## Materials and Methods

### Dataset

The first step involved defining the criteria for content selection. This step was crucial for ensuring that the subsequent stages were aligned with the core objectives and standards of diabetes self-management education. The foundation for content selection was the National Standards for Diabetes Self-Management Education.<sup>18</sup>



**Figure 2.** Article selection inclusion and exclusion criteria for dataset.

These standards propose a multifaceted framework for diabetes education, emphasizing the following critical aspects of diabetes education: integration of nutritional management within lifestyle parameters, safe and effective medication use, and strategies for the prevention, detection, and treatment of acute complications. We also included keywords related to DFUs, such as signs of infection, wound healing, and cold feet. A comprehensive list of 58 keywords and phrases was generated by physicians in the next step using the criteria established in the previous step. This list served as the foundation for the subsequent literature search, ensuring that the scope of the search was both relevant and exhaustive. See Table 1 for example keywords and phrases.

### Literature Search, Query Generation

Three rounds of literature search, query generation, and validations were conducted. Each round aimed to address knowledge disparities highlighted during its previous round of validation (Figure 2). During each round of literature search, the expert team of research editors ensured that each article was subject to intensive criteria: accuracy of the information, the credibility of references, adherence and relevance to the National Standards for Diabetes Self-Management

Education, and publication date post-2010. The editor team identified patient-centered resources by keeping in mind the information could be understood at an eighth-grade literacy level. One to three articles were found per keyword. One reviewer screened each article for inclusion. Questions related to DFUs were suggested by Podiatrists who work with these subsets of patients. The team focused on gathering literature from reputable sources such as the Centers for Disease Control and Prevention (CDC), the American Heart Association (AHA), the American Diabetes Association (ADA), and the International Working Group on the Diabetic Foot (IWGDF).

### Model Design and Development

The authors used Python as the language programming and database technologies, integrated with OpenAI's API as the supportive LLM to implement the RAG model. In this model, the LLM is not used to search and find the answers to a question, as it is used in applications such as ChatGPT. The selected documents are used as the context and reference point to find the closest answer to the query submitted by the user and provide a conversational chat experience by leveraging the LLM capabilities. In producing the answers, both zero-shot and few-shot learning were employed, whereas zero-shot learning enables a model to recognize and classify objects it has never seen during training by leveraging auxiliary information or attributes that describe each class. Few-shot learning, by contrast, involves training a model with a very small amount of labeled data, such as only a few examples of each category. We used few-shot learning to train the model on the desired format of answers and their attributes, such as the length of the answers and language simplicity.

### Validation

The evaluators created a set of sample questions to evaluate the model's answers. Each round of questions was designed to simulate various scenarios and challenges typically encountered in diabetes self-management. The five medically expert reviewers were trained to produce the queries that a patient would ask, review the model's answers, provide ground truth, and judge the correctness and completeness of the answers for each query. The sample questions were developed based on the keywords established earlier. Editors drafted the questions by establishing commonly asked patient questions related to each keyword in addition to considering the various stages of diabetes. The content of the questions were also based on publicly available documents such as NIH and CDC. After in-depth reading of these documents, the editor team created questions that the document should be able to answer for a patient. The environment was web-based, designed to ensure maximum usability and avoid cognitive overload. The AI model was then employed

**Table 1.** Example Keywords and Phrases.

National standards for diabetes self-management education	Example keywords and phrases
Describing the diabetes disease process and treatment options	1. Diabetes pathophysiology 2. Glycemic control
Incorporating nutritional management into lifestyle	1. Nutritional therapy for diabetes 2. Carbohydrate counting
Incorporating physical activity into lifestyle	1. Physical activity/exercise for diabetics 2. Incorporating exercise into daily routine
Using medication(s) safely and for maximum therapeutic effectiveness	1. Diabetes drug therapy 2. Oral hypoglycemic agents
Monitoring blood glucose and other parameters and interpreting and using the results for self-management decision making	1. Self-monitoring of blood glucose (SMBG) 2. Continuous glucose monitoring (CGM)
Preventing, detecting, and treating acute complications	1. Diabetic emergencies 2. Hypoglycemia management
Preventing, detecting, and treating chronic complications	1. Cardiovascular complications 2. Wound care in diabetes
Developing personal strategies to address psychosocial issues and concerns	1. Stress management in diabetes 2. Diabetes support networks
Developing personal strategies to promote health and behavior change	1. Lifestyle modification for diabetics 2. Setting goals for diabetes management

to formulate answers to the sample questions. This step was critical in evaluating the model's ability to provide accurate, relevant, and contextually appropriate responses based on the compiled knowledge.

The first round of review involved laying out a foundation of knowledge for the AI model. Using the comprehensive list of keywords, the expert research editors carefully selected approximately 200+ documents from reputable sources. Evaluated by medical professionals, 40 documents sourced from the repository of materials were also included in the first AI model. Five questions per keyword were generated from the utilization of keywords provided by physicians. The team diligently posed 175 questions to the AI platform and validated its responses. For sample questions and answers that failed or passed, see Table 1.

The second round of review emphasized patient-centered knowledge of diabetic management. The research editing team compiled a wide variety of 54 self-care resources, infographics, and articles that were digestible, explanatory, and legitimized by medical professionals. The second round, employing the same set of 175 questions from the first phase and an additional 24 questions focusing on DFUs, was conducted to evaluate the system's improvement with an expanded knowledge base.

The third round of literature search consisted of bridging knowledge gaps with supplemental information about more nuanced diabetic foot care topics. Topics were chosen based on validation processes that highlighted insufficient information to produce qualified answers. Evaluators entered each query and, upon seeing the answer, had the option to press a thumbs-up icon for "Passed" or a thumbs-down icon for "Failed." Subsequently, evaluators could enter text feedback for their validation. A Passed answer

did not require details; however, we used a method that scrutinized the model's responses across multiple dimensions.

The evaluation was structured around five critical criteria inspired by the criteria that Google's researchers introduced to validate the performance of Med-PaLM, Google's LLM, for medical knowledge extraction,<sup>3</sup> as follows: (1) Factuality: to ensure the model's outputs are accurate; (2) Comprehension: to evaluate the model's ability to explain the answers for a person with an eighth-grade literacy level and avoid medical jargon; (3) Reasoning: to assess the model's capability to synthesize information logically; (4) Possible Harm: to identify and mitigate potential risks or adverse effects stemming from the model's answers; and (5) Bias: to detect and address any inclination of the model to favor or discriminate against specific groups or individuals based on age, gender, ethnicity, or socioeconomic status. Information bias was a category that the editor team heavily focused on when posing questions and reading answers based on the criteria mentioned for the "Bias" category.

In addition, based on the National Standards, we created 23 topics with true or false phrases to test the model further to measure parameters such as precision, recall, and F1 score. Some topic examples are Type I and II diabetes, Diabetes and Exercise, Diabetes and Mental Health, Diabetes Medications, Technology and Diabetes-Self Management, Falls in Diabetics, Ketoacidosis, Hypoglycemia and Hyperglycemia, Diabetes and Immune System Health, Chronic Diabetes Complications—Nephropathy and Neuropathy, Chronic Diabetes Complications, Diabetes and Eye Care, Diabetic Foot Wound Prevention, Managing Your Diabetic Foot Wound, Diabetic Foot Wound Infections, Foot Deformities, Limb Preservation, and Offloading.

**Table 2.** Question and Answer Example.

Question	AI answer	Evaluation
What signs should I watch for if my blood sugar is too high or too low?	<p>If your blood sugar gets too high, a condition also known as hyperglycemia, you might notice signs like feeling especially thirsty, going to the bathroom more often, feeling tired and hungry, and having blurry vision. Sometimes, you might not feel any change at all, even if your sugar is high.</p> <p>On the other hand, if your blood sugar drops too low, which is called hypoglycemia, you might feel shaky, sweaty, dizzy, weak, or hungry. You could also feel nervous, grouchy, or have a fast heartbeat, impaired vision, fatigue, a headache, or become irritable.</p> <p>Remember, everyone's body reacts a bit differently, and you might not have all of these symptoms. It's essential to regularly check your blood sugar levels to stay informed about your condition. And as always, it's very important to talk to your health care provider if you're experiencing high or low blood sugar often or if you have other unexplained symptoms.</p>	Passed. Meets all criteria for evaluation.
Do I have to cut out all sugar to manage my diabetes?	No, you don't have to eliminate all sugar from your diet to manage your type 1 diabetes effectively.	Failed. The question does not specifically ask about Type 1 diabetes, but the answer is specific to it.

**Table 3.** Question and Answer Formulation.

Type	Model responses		
	Question	AI answer	AI reason
True positive	C-peptide testing is used to assess insulin production.	True	Yes, it is correct. C-peptide testing is used to understand how much insulin a person's body is making, as mentioned in the context.
True negative	Beta cells in the pancreas produce glucagon.	False	Beta cells in the pancreas produce insulin, not glucagon. It's the alpha cells that produce glucagon.
False positive	Gestational diabetes only occurs during pregnancy.	True	Gestational diabetes is defined as glucose intolerance that begins or is first recognized during pregnancy.
False negative	Stress can influence blood sugar levels by directly suppressing the immune system.	False	The statement is false because while stress can indeed affect blood sugar levels in people with diabetes, it does so by triggering the release of hormones that can interfere with insulin effectiveness, not by directly suppressing the immune system.

## Results

Our results revealed that the model successfully extracted relevant information for commonly asked questions by a layperson.

We deployed Zero-shot and Few-Shot learning prompts to evaluate the model. Zero-shot learning showed low performance with little content to provide the correct answers. By increasing the content, the number of answers that were not factually correct became zero; however, the number of answers that failed due to comprehension doubled. Both problems were resolved by implementing few-shot learning with the same amount of data in the knowledge base (see Table 2).

In the second validation phase, we designed a method to test the Accuracy, Precision, Recall, Specificity, and F1-score for the model. A supplementary examination was carried out,

in which the model encountered a set of True or False statements and was subsequently tasked with identifying their truthiness while also providing a reason. The statements taken from medically approved references had a balanced distribution of True and False. Each sentence was manually labeled as True or False with an evaluator, addressing various aspects of diabetes (Table 3).

Comparing the human evaluator's labels and AI's answers revealed a discrepancy between the results. Upon analyzing the model's reasonings, it became apparent that some statements were initially mislabeled by the human evaluators due to vague and ambiguous questions. Consequently, these statements were clarified and relabeled, and the model's scores were re-evaluated, resulting in the adjusted scores in (Table 4). Notably, this adjustment led to higher scores compared to the initial evaluation. Initially, the model scored between 84% and 89% across the five measurements.

**Table 4.** Phase I Results.

Validation criteria	Validation rounds		
	Round 1	Round 2	Round 3
Correct answers	74.86%	71.36%	94%
Factuality errors	45.45%	0%	0%
Comprehension error	38.64%	78.95%	0%

**Table 5.** Phase II Results.

Validation criteria	Validation rounds	
	Initial	Adjusted
Accuracy	0.89	0.98
Precision	0.94	0.99
Recall	0.84	0.98
Specificity	0.95	0.99
F1-score	0.89	0.98

After adjusting the answers, the model achieved a 98% Accuracy, Recall, and F1-score. A score of 99% was obtained for specificity and precision (Table 5).

## Discussion

In this study, we established and evaluated a RAG model, using innovative prompt engineering techniques to extract medical knowledge about diabetes and diabetic foot care for laypersons with an eighth-grade literacy level. The model underwent several rounds of validation, focusing on content quantity and prompt engineering, to ensure that its performance aligned with the desired criteria.

## Limitations

Our findings indicate that while innovative prompt engineering techniques and a comprehensive dataset can enhance answer quality, residual gaps may persist. These gaps necessitate further validation through reinforcement learning, augmented with feedback from human reviewers or AI agents.

Moreover, due to a lack of knowledge about the person who is asking the questions, our validation process did not show sensitivity to Information Bias, a critical factor in health, particularly in precision health domains where scientific results may skew the participant population profile.

## Future Work

Several avenues warrant future research from the limitations listed above. First, a deeper exploration of the optimal

content for RAG models to generate comprehensive answers is needed. Second, the impact of linguistic nuances on the model's output requires investigation. In this study, we utilized the ChatGPT-4 model for its conversational output capabilities; however, variations across different language models could influence the nature and effectiveness of prompts and instructions, affecting the resultant answers. Third, the efficacy of few-shot learning in improving model comprehension is another vital research domain. Specifically, determining the optimal number and format of examples that significantly bolster the model's understanding is crucial. Finally, due to the lack of information per user mentioned in the limitations section, integration of wearable or stationary remote monitoring devices will be crucial to personalize the answers. Artificial intelligence education systems and remote monitoring devices that view the individual as a full partner in their health care empower patients regarding their own health care.

## Conclusion

Improving health literacy through patient education is paramount to addressing the problem of limb loss in the diabetic population. In addition to affecting patient well-being through improved outcomes, improved physician well-being is an essential outcome of a self-management model for patient health education. We tested a RAG model for knowledge extraction in diabetes and DFUs with respect to vetted content, according to the National Standards for Diabetes Self-Management Education. The answers provided by the model were subjected to a rigorous validation process and involved an expert review and comparison with established guidelines and literature. Through several rounds of modifications and validations, we tested the relationship between the amount of content provided to the model and implementing innovative prompt engineering to provide accurate answers. In addition to the factuality of the answers, we included four additional criteria, comprehension, reasoning, information bias, and possible harm, to evaluate the answers. The study confirmed that the model achieved a 98% score for Accuracy, Recall, and F1-score with the appropriate amount of content and utilizing few-shot prompting. A score of 99% was obtained for the specificity and precision. Furthermore, we investigated discrepancies between the AI answers and human evaluators' answers; on seven out of ten occasions, the AI's answer was correct. In addition, the model self-identified when it did not have sufficient data in its knowledge base to find a proper answer to a question.

The study demonstrated that with appropriate content volume and few-shot learning prompts, the RAG model achieved 98% accuracy, confirming its capability to offer user-friendly and understandable medical information for a layperson with an eighth-grade literacy level.

## Abbreviations

ADA, American Diabetes Association; AHA, American Heart Association; CDC, Centers for Disease Control and Prevention; GAI, general artificial intelligence; IWGDF, International Working Group on the Diabetic Foot; NIH, National Institute of Health; RAG, retriever-augmented generation.

## Acknowledgments

The authors would like to thank Kevin Chyung and Issa Meinzer for contributing to the article search and validating the answers. The authors would also like to thank Joseph Mottahedeh and Ali Bolourian for their contribution to programming the model and prompt engineering discussions.

## Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Shayan Mashatian is an officer and shareholder, Shereen Aziz is an employee, and Ilia Alenabi is a research student at Silverberry Group, Inc.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was sponsored by Silverberry Group, Inc. This study is partially supported by the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases Award Number 1R01124789-01A1. This study is partially supported by the National Science Foundation (NSF) Center to Stream Healthcare in Place (#C2SHiP) CNS Award Number 2052578.

## ORCID iD

Shayan Mashatian  <https://orcid.org/0000-0003-2131-3367>

## References

1. Qiu Y, Jin Y. ChatGPT and finetuned BERT: a comparative study for developing intelligent design support systems. *Intell Syst Appl*. 2024;21:200308. doi:10.1016/j.iswa.2023.200308.
2. Bagde H, Dhopte A, Alam MK, Basri R. A systematic review and meta-analysis on ChatGPT and its utilization in medical and dental research. *Heliyon*. 2023;9(12):e23050. doi:10.1016/j.heliyon.2023.e23050.
3. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. *Nature*. 2023;620:172-180. doi:10.1038/s41586-023-06291-2.
4. Lweis P. Retrieval-augmented generation for knowledge-intensive NLP tasks. *arXiv:2005.11401v4 [cs.CL]*. 2021.
5. McDermott K, Fang M, Boulton AJM, Selvin E, Hicks CW. Etiology, epidemiology, and disparities in the burden of diabetic foot ulcers. *Diabetes Care*. 2023;46(1):209-221. doi:10.2337/dci22-0043.
6. Armstrong DG, Tan T-W, Boulton AJM, Bus SA. Diabetic foot ulcers: a review. *JAMA*. 2023;330:62-75. doi:10.1001/jama.2023.10578.
7. Armstrong DG, Swerdlow MA, Armstrong AA, Conte MS, Padula WV, Bus SA. Five year mortality and direct costs of care for people with diabetic foot complications are comparable to cancer. *J Foot Ankle Res*. 2020;13(1):16.
8. Hoffstad O, Mitra N, Walsh J, Margolis DJ. Diabetes, lower-extremity amputation, and death. *Diabetes Care*. 2015;38:1852-1857.
9. Meloni M, Izzo V, Giurato L, Lázaro-Martínez JL, Uccioli L. Prevalence, clinical aspects and outcomes in a large cohort of persons with diabetic foot disease: comparison between neuropathic and ischemic ulcers. *J Clin Med*. 2020;9:1780.
10. International Diabetes Federation. *The Diabetic Foot*. Brussels, Belgium: International Diabetes Federation; 2020. <https://www.idf.org/our-activities/care-prevention/diabetic-foot.html>. Accessed August 1, 2022.
11. Armstrong DG, Mills JL, Molina M, Molnar JA. *Nutrition Interventions in Adults With Diabetic Foot Ulcers*. Guideline Central; 2021. <https://www.guidelinecentral.com/guideline/502765/pocket-guide/502768/>. Accessed June 26, 2022.
12. Donnelly HR, Clarke ED, Collins CE, et al. Most individuals with diabetes-related foot ulceration do not meet dietary consensus guidelines for wound healing. *Int Wound J*. 2023;21:e14483. <https://onlinelibrary.wiley.com/doi/abs/10.1111/iwj.14483>. Accessed November 11, 2023.
13. Kessels RP. Patients' memory for medical information. *J R Soc Med*. 2003;96(5):219-222. doi:10.1177/014107680309600504.
14. Paterick TE, Patel N, Tajik AJ, Chandrasekaran K. Improving health outcomes through patient education and partnerships with patients. *Proc (Bayl Univ Med Cent)*. 2017;30(1):112-113. doi:10.1080/08998280.2017.11929552.
15. Thomas LR, Ripp JA, West CP. Charter on physician well-being. *JAMA*. 2018;319(15):1541-1542. doi:10.1001/jama.2018.1331.
16. Timmers T, Janssen L, Kool R, Kremer J. Educating patients by providing timely information using smartphone and tablet apps: systematic review. *J Med Internet Res*. 2020;22(4):e17342. <https://www.jmir.org/2020/4/e17342>. Accessed April 30, 2024.
17. American Medical Association. 6 big things that must change to beat physician burnout. <https://www.ama-assn.org/practice-management/physician-health/6-big-things-must-change-beat-physician-burnout>. Accessed April 30, 2024.
18. Funnell MM, Brown TL, Belinda P, et al. National standards for diabetes self-management education. *Diabetes Care*. 2010;33(suppl 1):S89-S96. doi:10.2337/dc10-S089.